

4F² EEPROM NROM MEMORY ARRAYS WITH VERTICAL DEVICES

Related Application

[0001] This application is related to U.S. Application Serial No. 10/769,116 (Attorney Docket No. 400.261US01) filed January 30, 2004, entitled "VERTICAL DEVICE 4F² EEPROM MEMORY," which is commonly assigned.

Technical Field of the Invention

[0002] The present invention relates generally to integrated circuits and in particular the present invention relates to EEPROM memory devices.

Background of the Invention

[0003] Memory devices are typically provided as internal storage areas in the computer. The term memory identifies data storage that comes in the form of integrated circuit chips. There are several different types of memory used in modern electronics, one common type is RAM (random-access memory). RAM is characteristically found in use as main memory in a computer environment. RAM refers to read and write memory; that is, you can both write data into RAM and read data from RAM. This is in contrast to read-only memory (ROM), which permits you only to read data. Most RAM is volatile, which means that it requires a steady flow of electricity to maintain its contents. As soon as the power is turned off, whatever data was in RAM is lost.

[0004] Computers almost always contain a small amount of ROM that holds instructions for starting up the computer. Unlike RAM, ROM cannot be written to. An EEPROM (electrically erasable programmable read-only memory) is a special type non-volatile ROM that can be erased by exposing it to an electrical charge. EEPROM comprise a large number of memory cells having electrically isolated gates (floating gates). Data is stored in the memory cells in the form of charge on the floating gates. Charge is transported to or removed from the floating gates by specialized programming and erase operations, respectively.

[0005] Yet another type of non-volatile memory is a Flash memory. A Flash memory is a type of EEPROM that is typically erased and reprogrammed in blocks instead of one byte at a time. A typical Flash memory comprises a memory array, which includes a large number of memory cells. Each of the memory cells includes a floating gate field-effect transistor capable of holding a charge. The data in a cell is determined by the presence or absence of the charge in the floating gate. The cells are usually grouped into sections called "erase blocks." The memory cells of a Flash memory array are typically arranged into a "NOR" architecture (each cell directly coupled to a bit line) or a "NAND" architecture (cells coupled into "strings" of cells, such that each cell is coupled indirectly to a bit line and requires activating the other cells of the string for access). Each of the cells within an erase block can be electrically programmed in a random basis by charging the floating gate. The charge can be removed from the floating gate by a block erase operation, wherein all floating gate memory cells in the erase block are erased in a single operation.

[0006] One recent type of Flash memory is a nitride read only memory (NROM). NROM has some of the characteristics of Flash memory but does not require the special fabrication processes of a conventional Flash memory, thus NROM integrated circuits can be implemented using a standard CMOS process. Because of their unique device characteristics, some NROM memory cells can also store multiple data bits in each cell (typically two bits each).

[0007] Figure 1 illustrates a cross-sectional view of a typical prior art NROM memory cell with a channel length, L. The cell is comprised of a control gate 100 formed on top of an oxide-nitride-oxide (ONO) layer. This layer is comprised of an oxide layer 101 on top of a nitride 103 layer upon which the charge is stored for the various states of the cell. In one embodiment, the cell has trapping areas 105 and 106 for storing two bits of data on the nitride layer 103. The nitride layer 103 is deposited on another oxide layer 104 that is on the substrate.

[0008] Two source/drain regions 109 and 111 are at either end of the gate 100. The source/drain regions 109 and 111 are connected by a channel area 110 between the two source/drain regions 109 and 111. The function of each source/drain region 109 or 111

(i.e., whether source or drain) depends upon which bit trapping area 105 or 106 is being read or written. For example, in a read operation, if the carrier is input at the left side source/drain region 111 and output from the right side region 109, the left side is the source 111 and the right side is the drain 109 and the data bit charge is stored on the nitride 103 at the source end 111 for bit trapping area 106.

[0009] As integrated circuit processing techniques improve, manufacturers try to reduce the feature sizes of the devices produced and thus increase the density of the IC circuits and memory arrays. In NROM memory arrays in particular, the channel length of the NROM memory cells that make up the memory array and spacing between memory cells in the strings have a large effect on the number of memory cells that can be placed in a given area and thus a direct impact on the density of the array and size of the resulting memory device. In addition, in many cases the feature sizes of the devices are limited by the device characteristics before the minimum feature size that the process is capable of is reached. In NROM devices in particular, as the channel length is reduced, a minimum size is typically reached that is primarily dictated by the device operational characteristics. Figure 2 illustrates a typical prior art planar NROM device that has a channel length that is less than 100 nm. In this case, the channel length is so short that the trapping areas 205 and 206 of the two data bits/cell NROM device overlap. This overlap may cause data write and read errors.

[0010] A further problem in floating gate and NROM memory cell arrays is the issue of overerased memory cells. An NROM memory cell is structurally similar to a MOSFET transistor, with a control gate separated from a channel, source, and drain by an insulator. In addition, embedded in the insulator is an isolated trapping layer. As in a MOSFET transistor, current flows when the NROM memory cell/transistor is selected or activated, charge trapped on the trapping layer affects the amount of current flow in the NROM transistor, depending on its direction of operation (current flow in the channel), effectively raising or lowering its threshold in the direction of operation. In programming or erasing an NROM memory cell, charge is transported to or from the electrically insulated trapping layer of the NROM transistor. If too much charge is removed from the trapping layer of the NROM transistor/memory cell it will flow

current in this direction of operation even when it is not selected. NROM transistors in this overerased state can affect current flow on shared bitlines and/or memory strings and thus potentially corrupt data read from other memory cells these common bitlines and/or memory strings.

[0011] For the reasons stated above, and for other reasons stated below which will become apparent to those skilled in the art upon reading and understanding the present specification, there is a need in the art for a method and architecture for producing a more closely spaced and, thus, higher density NROM memory array without trapping area overlap and with improved overerase handling properties.

Summary of the Invention

[0012] The above-mentioned problems with producing a more closely spaced and higher density NROM memory array with improved overerase handling properties and other problems are addressed by the present invention and will be understood by reading and studying the following specification.

[0013] NROM EEPROM memory devices and arrays, in accordance with embodiments of the present invention, facilitate the utilization of vertical NROM memory cells and select gates in NOR or NAND high density memory architectures. Memory embodiments of the present invention utilize vertical select gates and NROM memory cells to form NOR and NAND NROM architecture memory cell strings, segments, and arrays. These NROM memory cell architectures allow for an improved high density memory devices or arrays with integral select gates that can take advantage of the feature sizes semiconductor fabrication processes are generally capable of and yet do not suffer from charge separation issues in typical multi-bit NROM cells. The memory cell architectures also allow for mitigation of disturb and overerase issues by placing the NROM memory cells behind select gates that isolate the memory cells from their associated bit/data lines.

[0014] For one embodiment, the invention provides a memory device comprising a NOR architecture NROM memory array formed on a substrate having a plurality of pillars and associated intervening trenches. A plurality of memory cell structures, each

memory cell structure comprising, an NROM memory cell is formed vertically on a first sidewall of a trench. A select gate is formed on a second sidewall of the trench. The select gate is coupled to the NROM memory cell by a first source/drain region formed at the bottom of the trench.

[0015] For another embodiment, the invention provides an NROM memory cell structure comprising a substrate, having two raised areas, defining a trench therebetween. An NROM memory cell is formed vertically on a first sidewall of the trench, and a select gate memory cell is formed vertically on a second sidewall of the trench. The NROM memory cell is coupled to the select gate by source/drain regions at the bottom of the trench.

[0016] For yet another embodiment, the invention provides a method of forming an NROM memory cell structure comprising forming two raised areas on a substrate, the raised areas defining an associated intervening trench. The method further includes forming an NROM memory cell on a first sidewall of the trench, forming a select gate on a second sidewall of the trench, and forming a source/drain region at the bottom of the associated intervening trench.

[0017] For a further embodiment, the invention provides a NAND architecture NROM memory cell string comprising a substrate, comprising one or more raised areas, defining trenches therebetween. A plurality of NROM memory cells are formed vertically on the sidewalls of the trenches and are coupled in a serial string by source/drain regions formed at the top of the one or more raised areas and at the bottom of the one or more trenches. Additionally, a first NROM memory cell of the string is coupled to a first select gate.

[0018] For yet a further embodiment, the invention provides a memory array comprising a substrate having a plurality of pillars and associated intervening trenches. A plurality of NROM memory cells are formed vertically on the sidewalls of the plurality of pillars and trenches, wherein the plurality of NROM memory cells are coupled into a plurality of NAND architecture memory strings by source/drain regions formed at the top of the plurality of pillars and at the bottom of the associated trenches. Additionally, a first NROM memory cell of each NAND architecture memory string is

coupled to a first vertical select gate and a last NROM memory cell of each NAND architecture memory string is coupled to a second vertical select gate.

[0019] For another embodiment, the invention provides a method of forming a NAND architecture memory cell string comprising forming one or more raised areas on a substrate, the raised areas defining associated intervening trenches. The method further includes forming a plurality of NROM memory cells on the sidewalls of the one or more raised areas, forming one or more source/drain regions on the top of the one or more raised areas and at the bottom of the one or more associated intervening trenches, and forming a first vertical select gate coupled to a first NROM memory cell of the string and a second vertical select gate coupled to a last NROM memory cell of the string.

[0020] Other embodiments are also described and claimed.

Brief Description of the Drawings

[0021] Figure 1 shows a cross-sectional view of a typical prior art NROM cell.

[0022] Figure 2 shows a cross-sectional view of a typical prior art NROM cell with a channel less than 100 nm.

[0023] Figures 3A and 3B detail erase operations in NROM memory cells.

[0024] Figures 4A-4C detail a planar NAND NROM memory array of the prior art.

[0025] Figures 5A-5C detail a planar NOR NROM memory array of the prior art.

[0026] Figures 6A-6C detail vertical NOR NROM memory array and cells in accordance with embodiments of the present invention.

[0027] Figures 7A and 7B detail vertical NAND NROM memory array and cells in accordance with embodiments of the present invention.

[0028] Figures 8A-8C detail formation of a vertical NAND architecture NROM memory array having NROM memory cells and select gates in accordance with embodiments of the present invention.

[0029] Figure 9 details a block diagram of an electronic system in accordance with embodiments of the present invention.

Detailed Description of the Invention

[0030] In the following detailed description of the preferred embodiments, reference is made to the accompanying drawings that form a part hereof, and in which is shown by way of illustration specific preferred embodiments in which the inventions may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention, and it is to be understood that other embodiments may be utilized and that logical, mechanical and electrical changes may be made without departing from the spirit and scope of the present invention. The terms wafer and substrate used previously and in the following description include any base semiconductor structure. Both are to be understood as including silicon-on-sapphire (SOS) technology, silicon-on-insulator (SOI) technology, thin film transistor (TFT) technology, doped and undoped semiconductors, epitaxial layers of silicon supported by a base semiconductor, as well as other semiconductor structures well known to one skilled in the art. Furthermore, when reference is made to a wafer or substrate in the following description, previous process steps may have been utilized to form regions/junctions in the base semiconductor structure. The following detailed description is, therefore, not to be taken in a limiting sense, and the scope of the present invention is defined only by the claims and equivalents thereof.

[0031] NROM EEPROM memory devices and arrays, in accordance with embodiments of the present invention, facilitate the utilization of vertical NROM memory cells and select gates in NOR or NAND high density memory architectures. Memory embodiments of the present invention utilize vertical select gates and NROM memory cells to form NOR and NAND NROM architecture memory cell strings, segments, and arrays. These NROM memory cell architectures allow for improved high density memory devices or arrays with integral select gates that can take advantage of the feature sizes semiconductor fabrication processes are generally capable of and yet do not suffer from charge separation issues in typical multi-bit NROM cells. By constructing NROM memory cells/gates in a vertical orientation, embodiments of the

present invention allow for increases in memory array cell density and improved utilization of process minimum feature size capabilities while maintaining an appropriate NROM memory cell channel length to allow for effective charge separation. The memory cell architectures also allow for mitigation of disturb and overerase issues by placing the NROM memory cells behind select gates that isolate the memory cells from their associated bit/data lines and/or source lines.

[0032] As stated above, as integrated circuit processing techniques improve, manufacturers try to reduce the feature sizes of the devices produced and thus increase the density of the IC circuits and memory arrays. In many cases, the feature sizes of the devices are limited by the device characteristics before the minimum feature size that the process is capable of is reached. In both NAND and NOR architecture NROM EEPROM memory arrays, as the channel length is reduced and the spacing between memory cells in the strings are reduced, a minimum size is reached that is dictated by the operational characteristics of the NROM memory cell devices that make up the memory array. As a result, the maximum density of an array of memory cells is limited even if the process technology can attain even smaller features and/or channel lengths. In particular, this is an issue in higher capacity memory types where small changes in the memory cell footprint (e.g., memory cell channel widths) and the cell density in the array can have a large effect on the overall array size and the resulting storage capacity.

[0033] As stated above, programming an NROM memory cell involves the insertion and storage of charge on the memory cell's trapping layer. However, as also stated above, because of their unique device characteristics, NROM memory cells can also store multiple data bits in each cell, storing a data bit in the trapping layer of the device near each source/drain region for two bits in a typical NROM memory cell transistor. In this, an NROM memory cell transistor is read and programmed in both a forwards and backwards direction of current flow in the source/drain regions (interchanging their source/drain function) to allow access to programming, reading, and erasing the two stored data bits. The function of each source/drain region (i.e., whether source or drain) depends upon which bit trapping area is being read or written. For example, in a read operation, if the carrier is input at the left side source/drain region and output from the

right side region, the left side is the source and the right side is the drain and the data bit charge is stored on the nitride at the source end for bit trapping area. Because of the localized storage of electrons in the trapping layer, while reading the NROM memory cell only the charge stored in the trapping layer nearest the source/drain region operating as the source affects the current flow through the device. The charge stored near the other source/drain region is “read through” and has minimal influence. The bits in an NROM memory cell transistor are programmed in the reverse bias/current flow direction from the read direction for each stored data bit.

[0034] Typically, an NROM memory cell electrically operates as an enhancement type MOS transistor, requiring a positive voltage applied across the gate and channel to flow current. The presence, or lack thereof, of a trapped charge on the trapping layer of the NROM transistor/memory cell near the source/drain region biased as the drain alters the threshold voltage characteristics of this transistor in this bias direction and thus the amount of current the transistor will flow at a given gate channel voltage. The effective threshold voltage of the NROM transistor memory cell in the selected bias direction increases with storage of electrons on the trapping layer near the source/drain region operating as the source and decreases with the removal of stored charge. Reading NROM memory cells in EEPROM memory arrays is then accomplished by accessing a memory cell with selected read voltage levels on the control gate and data/bit line coupled to the source/drain region biased as the source. The stored data is then typically sensed from the amount of current the NROM memory cell flows through the coupled bit/data line coupled to the source/drain region biased as the drain.

[0035] In programming NROM memory cells in EEPROM memory arrays, electrons are typically transferred to the trapping layer of the memory cell by one of Fowler-Nordheim tunneling (FN-Tunneling) or channel hot electron injection (HEI). Other forms of programming NROM memory cells, such as, substrate enhanced hot electron injection (SEHE), are also known and utilized, but may not be appropriate for multi-bit operation. FN-Tunneling is typically accomplished by applying a positive control gate voltage on the NROM memory cell with respect to its channel to tunnel inject electrons to the trapping layer. Channel hot electron injection (HEI) is typically

accomplished by applying a positive voltage on the control gate and the source/drain region biased as the drain of the NROM memory cell and a low voltage or ground on the source/drain region biased as the source to inject electrons to the trapping layer. In many cases the programming voltages are iteratively pulsed and the memory cell read to check the programming process and more accurately program the NROM memory cell.

[0036] Erasure of the NROM memory cells of EEPROM memory arrays is typically accomplished by conventional tunneling or negative voltages applied to the control gate voltages with respect to the channel. Alternatively, other forms of erasure, such as, substrate enhanced band to band tunneling induced hot hole injection (SEBBHH) can also be used for NROM memory cell erasure, but again, may not be suitable for multi-bit operation. To ensure uniformity, in many cases the EEPROM memory programs all the NROM memory cells in the data segment to be erased before applying the voltages to erase the memory cells. As with programming, the erasure voltages are typically iteratively pulsed and the memory cells checked after each pulse to verify erasure and return of the NROM memory cells to an un-programmed threshold voltage state.

[0037] Unfortunately, it is possible, during an erasure process, for too much charge to be removed from the trapping layer or individual trapping areas of an NROM memory cell transistor placing it in to an “overerased” state. In such cases, enough charge is removed that the threshold voltage of the NROM memory cell transistor is altered so that it operates as a depletion mode device, requiring a negative control gate-channel voltage to be applied to shut off current flow in the selected bias direction. In this state, the NROM memory cell transistor will flow a current even when it has not been selected by the memory, unless a negative voltage is applied to the control gate with respect to the source. This “overerased” state and the resulting current flow in the selected bias direction when the overerased NROM memory cell is unselected can interfere with attempts to read the values of other NROM memory cells that share the common bit lines with it, corrupting the read data.

[0038] Complicating the issue of overerasure in NROM memory cells is that not all NROM memory cells erase (remove charge from their trapping layers or trapping areas)

at the same rate of speed. Typically, one or more “fast erasing” memory cells will erase more quickly than the others of the group of cells selected for erasure. To minimize the possibility of inadvertently overerasing this group of fast erasing NROM memory cells most NROM EEPROM and Flash memory devices/arrays typically utilize the complex and time consuming iterative erase voltage pulse-memory cell verify process noted above to erase its NROM memory cells.

[0039] Figure 3A details a chart of a distribution of threshold voltages 300 of NROM memory cells in a given array, showing distribution bell curves of the threshold voltages of NROM memory cells placed in a programmed 302 and erased 304 state in a given bias direction as may be typical of an NROM memory device. Figure 3A also details a number of erased cells that have been inadvertently placed in an overerased (depletion mode) 306 state. Figure 3B details a bit line 310 of a memory array 312 and the current flow 314 through an overerased NROM memory cell 316 that has not been selected/activated for reading.

[0040] An additional issue that can affect the memory cells of an EEPROM or Flash memory array is “disturb.” Disturb typically happens when the elevated voltages used to program or erase an NROM memory cell segment or erase block “disturb” the programmed values stored in other trapping areas of other NROM memory cells corrupting them and causing errors when they are later read. These inadvertently disturbed cells typically share common word lines, bit lines, or source lines with the memory cells that are being programmed or erased.

[0041] By constructing their NROM memory cells vertically and isolating them with select gates, embodiments of the present invention allow for increases in memory array cell density and improved utilization of process minimum feature size capabilities, while maintaining the size of the memory cell channel to allow for appropriate device operation. In addition, by increasing the available surface area for transistors and incorporating select gates, which are also constructed in a vertical manner, embodiments of the present invention mitigate issues with overerasure and disturb, allowing for greater reliability and faster programming and erasure.

[0042] As previously stated, the two common types of EEPROM and Flash memory array architectures are the “NAND” and “NOR” architectures, so called for the similarity each basic memory cell configuration has to the corresponding logic gate design. In the NOR array architecture, the NROM memory cells of the memory array are arranged in a matrix similar to RAM or ROM. The gates of each NROM memory cell of the array matrix are coupled by rows to word select lines (word lines) and their source/drain regions are coupled to column bit lines. The NOR architecture NROM memory array is accessed by a row decoder activating a row of NROM memory cells by selecting the word line coupled to their gates. The row of selected memory cells then place their stored data values on the column bit lines by flowing a differing current from the coupled source line to the coupled column bit lines depending on their programmed states. A column page of bit lines is selected and sensed, and individual data words are selected from the sensed data words from the column page and communicated from the memory.

[0043] A NAND array architecture also arranges its array of NROM memory cells in a matrix such that the gates of each NROM memory cell of the array are coupled by rows to word lines. However each memory cell is not directly coupled to a source line and a column bit line. Instead, the memory cells of the array are arranged together in strings, typically of 8, 16, 32, or more each, where the memory cells in the string are coupled together in series, source to drain, between a common source line and a column bit line. This allows a NAND array architecture to have a higher memory cell density than a comparable NOR array, but with the cost of a generally slower access rate and programming complexity.

[0044] A NAND architecture NROM memory array is accessed by a row decoder activating a row of NROM memory cells by selecting the word select line coupled to their gates. In addition, the word lines coupled to the gates of the unselected memory cells of each string are also driven. However, the unselected memory cells of each string are typically driven by a higher gate voltage so as to operate them as pass transistors and allowing them to pass current in a manner that is unrestricted by their stored data values. Current then flows from the source line to the column bit line

through each NROM memory cell of the series coupled string, restricted only by the memory cells of each string that are selected to be read. This places the current encoded stored data values of the row of selected memory cells on the column bit lines. A column page of bit lines is selected and sensed, and then individual data words are selected from the sensed data words from the column page and communicated from the memory.

[0045] Figures 4A, 4B, and 4C show a simplified planar NROM memory array of a EEPROM or Flash memory device of the prior art. Figure 4A details a top view of a planar NOR architecture NROM memory array 400, a side view of the planar NROM memory array 400 is detailed in Figure 4B. In Figures 4A, 4B, and 4C, NROM memory cells 402 are coupled together in a NOR architecture memory array having bit lines 412 and word lines 406. The bit/source lines 412 are formed locally from N+ doped regions deposited in the substrate 408. Each NROM memory cell 402 has a gate-insulator stack formed between the N+ doped regions of a bit line 412, utilizing the N+ doped regions as a drain and source respectively. The gate-insulator stack is made of an insulator on top of a substrate 408, a trapping layer formed on the insulator, an intergate insulator formed over the trapping layer, and a control gate 406 (typically formed integral to the word line 406, also known as a control gate line) formed over the intergate insulator. Figure 4C details an equivalent circuit schematic 420 of the NOR architecture NROM memory array 400, showing NROM memory cells 402 coupled to the bit lines, word lines, and substrate connection 422.

[0046] Figures 5A, 5B, and 5C show a simplified planar NAND NROM memory array of a EEPROM or Flash memory device of the prior art. Figure 5A details a top view of a planar NAND NROM memory string 504 of a NAND architecture NROM memory array 500, a side view of the planar NAND NROM memory string 504 is detailed in Figure 5B. In Figures 5A and 5B, a series of NROM memory cells 502 are coupled together in a series NAND string 504 (typically of 8, 16, 32, or more cells). Each NROM memory cell 502 has a gate-insulator stack that is made of an insulator on top of a substrate 508, a trapping layer formed on the insulator, an intergate insulator formed over the trapping layer, and a control gate 506 (typically formed in a control

gate line, also known as a word line) formed over the intergate insulator. N+ doped regions are formed between each gate insulator stack to form the source/drain regions of the adjacent NROM memory cells, which additionally operate as connectors to couple the cells of the NAND string 504 together. Select gates 510, that are coupled to gate select lines, are formed at either end of the NAND NROM string 504 and selectively couple opposite ends of the NAND NROM string 504 to a bit line contacts 512. Figure 5C details an equivalent circuit schematic 520 of the NAND architecture NROM memory string 504, showing NROM memory cells 502 and substrate connection 522.

[0047] Embodiments of the present invention utilize vertical NROM memory cells and vertical gate structures. Methods of forming vertical memory cells are detailed in United States Patent Application No. 10/177,208, titled “Vertical NROM having a storage density of 1 bit per $1F^2$ ”, filed June 21, 2002, and United States Patent No. 5,936,274, titled “High density flash memory”, issued August 10, 1999, which are commonly assigned. Methods of forming vertical split control gates are detailed United States Patent No. 6,150,687, titled “Memory cell having a vertical transistor with buried source/drain and dual gates”, issued November 21, 2000, and United States Patent No. 6,072,209, titled “Four F^2 folded bit line DRAM cell structure having buried bit and word lines”, issued June 6, 2000, which are also commonly assigned.

[0048] Figures 6A-6C detail vertical NROM memory cells and select gates for a NOR architecture NROM memory array in accordance with embodiments of the present invention. Figure 6A details a side view of a simplified vertical NOR architecture memory cell structure 604, a side view of a vertical NOR memory array 600 is detailed in Figure 6B. Figure 6C details an equivalent circuit schematic 620 of the vertical NOR architecture NROM memory array 600, showing NROM memory cells 602 and substrate connection 622. It is noted that the NOR architecture NROM memory array 600 can be utilized in both EEPROM or Flash memories arrays and devices. As can be seen from Figures 6A and 6B, in a single vertical NOR architecture memory cell structure 604, two vertically formed transistors occupy the area that a single planar NROM transistor would occupy (an area of $4F^2$ squared when viewed from above, each

transistor having an area of $2F$ squared). Where “F” is the minimum resolvable photolithographic dimension in the particular process technology.

[0049] In Figure 6A, a vertically formed NROM memory cell 602 and select gate 610 are coupled together in series in a vertical NOR architecture memory cell structure 604. In creating the vertical NOR architecture memory cell structure 604 a trench 630 is formed in a substrate 608. The vertical NROM memory cell 602 and select gate 610 are then formed on the sidewalls of the trench 630. The vertical NROM memory cell 602 has a gate-insulator stack made of an insulator 620 formed on the surface of the sidewall, a trapping layer 622 (typically of nitride) formed on the insulator 620, an intergate insulator 624 formed over the NROM 622, and a control gate 606 (typically formed in a control gate line, also known as a word line) formed over the intergate insulator 624. In one embodiment, the substrate trench 630 is formed by patterning a masking material that is layered over the substrate 608 and anisotropically etching the trenches 630. The gate-insulator stack of the NROM memory cell 602 is formed in one embodiment by successive layering of each of the materials of the gate insulator stack over the trench 630, followed by a mask and directional etch of the deposit of each layer to leave only the material deposited on the sidewall of the trench 630. In another embodiment, differing layers of the gate-insulator stack are formed and then masked and directionally etched in a single step.

[0050] The vertical select gate 610 has a gate-insulator stack made of an insulator 642 formed on the opposite sidewall of the trench 630 with a control gate 644 formed over the insulator 642. The gate-insulator stack of the select gate 610 is formed by successive layering of each of the materials of the gate insulator stack over the trench 630, as with the vertical NROM memory cell 602, but skips the depositing of the insulator 620 and trapping layer 622 layers. In one embodiment of the present invention the insulator 620 and the trapping layer 622 of the NROM memory cell 602 are formed and then the intergate insulator 624/control gate 606 of the NROM memory cell 602 and the insulator 642/control gate 644 of the select gate 610 are formed consecutively.

[0051] N+ doped regions 626 are formed at the raised areas at the top and at the bottom of the trenches 630 to form the source/drain regions for the vertical NROM memory cell/gate-insulator stack 602 and select gate 610. The N+ regions also couple the memory cell 602 and select gate 610 together to form the vertical NOR memory structure 604 and additionally couple the vertical NOR architecture memory cell structure 604 to the bit lines 612. It is noted that the N+ source/drain regions 626 may be formed before or after the formation of the NROM memory cell 602 and select gate 610 gate-insulator stacks.

[0052] In Figure 6B, a vertical NOR architecture NROM memory array 600 is formed from a series of vertical NOR architecture memory cell structures 604. Each vertical NOR architecture memory cell structures 604 having a vertically formed NROM memory cell 602 and a coupled select gate 610, wherein the drain of the select gate is coupled to a first bit/data line 612 and a source/drain of the NROM 602 is coupled to a second bit/data line 612.

[0053] In creating the vertical NOR architecture NROM memory array 600 a series of substrate pillars 628 are formed in a substrate 608 with trenches 630 located between them. The vertical NROM memory cells 602 and select gates 610 are then formed on the sidewalls of the pillars 628 within the trenches 630 to form the vertical NOR architecture memory cell structures 604. The vertical NROM memory cells 602 and select gates 610 are formed in an alternating pattern (NROM-select gate, select gate-NROM, NROM-select gate, etc.) such that each pillar 628 has either select gates 610 or NROM memory cells formed on its sidewalls.

[0054] N+ doped regions 626 are formed at the top of the pillars 628 and at the bottom of the trenches 630 to form the source and drain regions. The N+ regions at the bottom of the trenches 630 couple the memory cell 602 and select gate 610 of each vertical NOR architecture memory cell structure 604 together. The N+ regions at the tops of the pillars 628 couple the drain of the select gate 610 and the source of the NROM memory cell 602 of each vertical NOR architecture memory cell structure 604 to the bit lines 612 respectively. It is again noted that the N+ source/drain regions 626

may be formed before or after the formation of the NROM memory cell 602 and select gate 610 gate-insulator stacks.

[0055] It is also noted that isolation regions, typically formed of an oxide insulator, can be used between adjacent rows of vertical NOR architecture memory cell structures 604 to isolate each row from its neighbors. These isolation regions can be extended into the substrate 608 to allow the formation of P-wells, where each P-well contains a single row of vertical NOR architecture memory cell structures 604 that can be biased in isolation from the other rows of the array 600. It is also noted that the control gate/word address lines 606 and select lines 640 can cross these isolation regions so that each control gate/word address line 606 and select line 640 controls the operation of multiple NROM memory cells 602 and select gates 610, respectively, across multiple rows of vertical NOR architecture memory cell structures 604.

[0056] As stated above, Figure 6C details an equivalent circuit schematic 620 of the vertical NOR architecture NROM memory array 600, showing NROM memory cells 602 and substrate connection 622. The vertical NROM memory cells 602 and select gates 610 are formed in an alternating pattern (NROM-select gate, select gate-NROM, NROM-select gate, etc.) so that the drain of the select gate 610 and the source/drain of the NROM memory cell 602 of adjacent vertical NOR architecture memory cell structures 604 are both coupled to a single bit line 612. It is noted that non-alternating patterns are also possible. It is noted that other array patterns of vertical NROM memory cells 602, select gates 610, and vertical NROM memory cell structures 604 are possible and will be apparent to those skilled in the art with the benefit of the present disclosure.

[0057] In the vertical NOR architecture NROM memory array 600 of Figures 6A-6C, the channel length of each NROM memory cell 602 and select gate 610 in a vertical NOR architecture memory cell structure 604 is determined by the depth of the trenches 630 and not by the minimum feature size. Due to the vertical form of the vertical NOR architecture NROM memory array 600 and vertical NOR architecture memory cell structures 604 of embodiments of the present invention, a NOR architecture NROM memory array can be produced that contain a vertical NROM memory cell 602 that

does not suffer from charge separation issues and a coupled select gate 610 in the space that would be utilized by a single conventional planar NROM memory cell.

[0058] The addition of a select gate 610 coupled between the bit line 612 and a source/drain of each NROM memory cell 602 allows for the NROM memory cell 602 to be isolated from the bit lines 612 and thus has advantages in both programming and erasing the NROM memory cells 602 of the vertical NOR architecture NROM memory array 600. In erasing, the coupled select gate 610 allows for avoidance of overerase issues with NROM memory cells 602 by isolating each memory cell 602 behind a select gate 610 so that, even if an NROM memory cell 602 is overerased into depletion mode, it will not corrupt the reading of other memory cells 602 on its coupled bit lines 612 by flowing current. As the possibility of corruption of data reads due to overerase is mitigated, this allows for the NOR architecture NROM memory array 600 to speed up its erasure processes by using larger erase pulses (in time or voltage) or even utilizing only a single erase pulse. In addition, by allowing the increase of erasure pulse time and voltage, the number of erasure verifications required are reduced or even eliminated. The coupled select gate 610 also allows for the isolated erasure of one or more NROM memory cells 602 allowing erasure of individual NROM memory cells 602, one or more selected data words, data segments, or erase blocks.

[0059] In programming, the coupled select gate 610 allows for mitigation of programming disturb of NROM memory cells 602 in the array 600 by the select gate 610 isolating or buffering the memory cells 602 from the bit lines 612. This allows for longer and higher voltage programming pulses to be used without increasing the possibility of disturb issues.

[0060] Figures 7A-7D detail vertical NROM cells, vertical select gates, and NAND architecture NROM memory strings in accordance with embodiments of the present invention. Figure 7A details a side view of a simplified vertical NAND architecture NROM memory string 704 with vertical select gates 710. Figure 7B details an equivalent circuit schematic 720 of the vertical NAND architecture NROM memory string 704, showing NROM memory cells 702 and substrate connection 734. Figure 7C details a side view of a simplified vertical NAND architecture NROM memory array

700 and Figure 7D details an equivalent circuit schematic of one embodiment of the present invention. Again, as can be seen from Figures 7A and 7C, in a vertical NAND architecture memory string 704, two vertically formed transistors occupy the area that each planar transistor would occupy (an area of $4F$ squared when viewed from above, each transistor having an area of $2F$ squared). Where “F” is the minimum resolvable photolithographic dimension in the particular process technology. Since each transistor can store a single bit of data the data storage density is one bit for each $2F$ squared unit area. Thus, for example, if $F=0.1$ micron then the storage density is 0.5 Giga bit per square centimeter.

[0061] In Figure 7A, a series of vertically formed NROM memory cells 702 are coupled together in a series NROM NAND string 704 (typically of 8, 16, 32, or more cells). In the vertical NAND NROM memory array string 704 of Figures 7A-7D, a series of substrate pillars 728 are formed in a substrate 708 with trenches 730 located between them. The vertical NROM memory cells 702 are then formed on the sidewalls of the pillars 728 within the trenches 730. Each vertical NROM memory cell 702 is formed on the sidewalls of the substrate pillars 728 (for two NROM memory cells 702 per trench 730) and has a gate-insulator stack made of an insulator 720 formed on the surface of the sidewall, a trapping layer 722 (typically of polysilicon) formed on the insulator 720, an intergate insulator 724 formed over the trapping layer 722, and a control gate 706 (typically formed in a control gate line, also known as a word line) formed over the intergate insulator 724.

[0062] In one embodiment the substrate pillars 728 and trenches 730 are formed by patterning a masking material that is layered over the substrate 708 and anisotropically etching the trenches 730. The gate-insulator stack of each NROM memory cell 702 are formed in one embodiment by successive layering of each of the materials of the gate insulator stack over the pillars 728 and trenches 730, followed by a mask and directional etch of the deposit of each layer to leave only the material deposited on the sidewall of the pillars 728. In another embodiment, differing layers of the gate-insulator stack are formed and then masked and directionally etched in a single step.

[0063] N+ doped regions 726 are formed at the top of the substrate pillars 728 and at the bottom of the trenches 730 between each vertical NROM memory cell/gate-insulator stack 702 to form the source and drain regions of the adjacent NROM memory cells 702 and couple the memory cells 702 together to form the vertical NAND architecture memory string 704. It is noted that the N+ source/drain regions 726 may be formed before or after the formation of the NROM memory cells/gate-insulator stack 702.

[0064] Select gates 710, that are coupled to gate select lines, are formed at either end of the NAND NROM memory string 704 and selectively couple opposite ends of the NAND NROM memory string 704 to a bit line contacts 712. The vertical select gates 710 have a gate-insulator stack made of an insulator 742 formed on a sidewall with a control gate 744 formed over the insulator 742. The gate-insulator stack of the select gates 710 are formed by successive layering of each of the materials of the gate insulator stack over the pillars 728 and trenches 730, as with the vertical NROM memory cell 702, but skips the depositing of the insulator 720 and trapping layer 722 layers. The N+ regions 726 also couple the first and last memory cell 702 of the vertical NAND architecture NROM string 704 to the select gates 710 and additionally couple the vertical NAND architecture NROM string 704 to the bit lines 712.

[0065] As stated above, Figure 7B details an equivalent circuit schematic of the vertical NAND architecture NROM memory array 700, showing the vertical NROM memory cells 702, select gates 710, bit line 712 connections, and substrate connection 734, in accordance with embodiments of the present invention. As can be seen, the schematic provides the same equivalent circuit as that of a conventional planar NAND architecture NROM memory string.

[0066] In Figure 7C, a section of vertical NAND architecture NROM memory array 700 of one embodiment of the present invention is formed from a series of vertical NAND architecture NROM memory cell strings 704. In Figure 7C, each pair of adjacent vertical NAND architecture NROM memory cell strings 704 in the vertical NAND architecture NROM memory array 700 are coupled through vertical select gates

710 to a common bit line 712 by a N+ doped region 726 formed at the top of a pillar 728.

[0067] It is also noted that isolation regions, typically formed of an oxide insulator, can be used between vertical NAND architecture NROM memory cell strings 704 to isolate each string 704 from its neighbors. These isolation regions can be extended into the substrate 708 to allow the formation of P-wells, where each P-well contains a single vertical NAND architecture NROM memory cell string 704 that can be biased in isolation from the other strings or rows of the array 700. It is also noted that the control gate/word address lines 706 and select lines 740 can cross these isolation regions so that each control gate/word address line 706 and select line 740 controls the operation of NROM memory cells 702 and select gates 710 respectively across multiple rows of vertical NAND architecture NROM memory cell strings 704.

[0068] As stated above, Figure 7D details an equivalent circuit schematic of the vertical NAND architecture NROM memory array of Figure 7C, showing NROM memory cells 702 and adjacent sting 704 connection.

[0069] In the vertical NAND architecture NROM memory array 700 of Figures 7A-7D, the channel length of each NROM memory cell 702 and select gate 710 in a vertical NAND architecture memory string 704 is determined by the depth of the pillars 728 and trenches 730 and not by the minimum feature size. Due to the vertical form of the NAND architecture NROM memory array 700 and NAND architecture memory strings 704 of embodiments of the present invention, a vertical NAND architecture NROM memory array string 704 and select gates 710 can be produced that typically has twice the density for a given string horizontal run length than a corresponding planar NAND architecture NROM memory array string.

[0070] The addition of select gates 710 coupled between the bit lines 712 and the NROM memory cells 702 of the vertical NAND architecture memory string 704 allows for the NROM memory cells 702 of the vertical NAND architecture memory string 704 to be isolated from one or both of the bit lines 712 and thus has advantages in both programming and erasing the vertical NAND architecture memory string 704. In erasing, the coupled select gates 710 allows for avoidance of overerasure issues with

NROM memory cells 702 by isolating each vertical NAND architecture memory string 704 behind one or more select gates 710 so that even if the NROM memory cells 702 of the vertical NAND architecture memory string 704 are overerased into depletion mode operation they will not corrupt the reading of other memory cells 702 on other vertical NAND architecture memory strings 704 that are coupled to the same bit line 712. As the possibility of corruption of data reads due to overerasure is mitigated, this allows for the vertical NAND architecture NROM memory array 700 to speed up erasure processes by using larger erase pulses or even a single erase pulse and by reducing or eliminating erasure verification. The coupled select gates 710 also allow for the isolated erasure of one or more NROM memory cells 702 allowing erasure of individual NROM memory cells 702, individual vertical NAND architecture memory strings 704, one or more selected data words, or erase blocks. In programming, the coupled select gates 710 allow for mitigation of programming disturb of NROM memory cells 702 in the array 700 by the select gates 710 isolating the memory cells 702 of the vertical NAND architecture memory strings 704 from the bit lines 712. This allows for longer and higher voltage programming pulses to be used without an increased issue with of disturb problems.

[0071] Figures 8A-8C detail three dimensional views of vertical NROM cells 802 and vertical select gates 810 of a vertical NAND NROM memory array 800 in accordance with embodiments of the present invention at several mid-fabrication stages. It is noted that a formation process that is similar to the NAND process may be utilized for formation of a vertical NOR NROM memory array 600 with the exception of placement and number of the select gates 810 and bit line and source line contacts. As stated above, in creating the vertical NAND NROM memory array 800, a series of substrate pillars 828 are formed in a substrate 808 with trenches 830 located between them. The vertical NROM memory cells 802 and select gates 810 are then formed on the sidewalls of the pillars 828 within the trenches 830. Between successive rows of substrate pillars 828, isolation regions 832 have been formed on the faces of the pillars 828 that are not utilized to form NROM memory cells 802 or select gates 810 to isolate each row of vertical NAND NROM memory strings 804 from the neighboring rows. These isolation regions 832 are typically formed of an oxide insulator.

[0072] As stated above, in creating each NROM memory cell gate-insulator stack 802, an insulator 820 is formed on the surface of the sidewall, a trapping layer 822 is formed on the insulator 820, an intergate insulator 824 is formed over the trapping layer 822, and a control gate 806 is formed over the intergate insulator 824. In creating each select gate gate-insulator stack 810, an insulator 842 is formed on the surface of the sidewall, and a control gate 844 is formed over the insulator 842.

[0073] In Figure 8A, the trenches 830 have been already formed by masking and anisotropically/directionally etching the trenches 830 in the substrate 808. N+ doped regions 826 have been formed at the top of the unformed substrate pillars 828 and at the bottom of the trenches 830 to form the source/drain regions of the NROM memory cells 802. The gate-insulator stack of each NROM memory cell 802 have been partially formed on the sidewalls of the trenches 830. In each trench 830, with the exception of the pillar 828 on which the select gates 810 are to be formed, are formed the insulator 820, the trapping layer 822, and the intergate insulator 824, by successive depositing, masking, and directional etching of layers of material.

[0074] In Figure 8B, the pillars 828 are formed and the space between each pillar in successive rows of NAND architecture NROM memory strings 804 are filled with an oxide to form isolation regions 832. In forming the pillars 828, the rows of NAND architecture NROM memory strings 804 are masked and directionally etched. This masking and etching process also divides the NROM layer into individual trapping layers 822.

[0075] In Figure 8C, the control gates/word lines 806 and select gates 810/select lines 840 are formed. In forming the control gates/word lines 806 and select gates 810/select lines 840, successive layers of insulator and polysilicon which will form the control gates/word lines 806 and select gates 810/select lines 840 is deposited over the pillars 828, trenches 830, and partially formed gate-insulator stacks of the NROM memory cells 802 of the memory array 800. A layer of masking material is then formed over the polysilicon layer and patterned. The excess masking material is removed and the memory array 800 is anisotropically/directionally etched to remove the undesired

portions of the deposited polysilicon and form the control gates/word lines 806 and select gates 810/select lines 840 on the sidewalls of the pillars 828 and trenches 830.

[0076] It is noted that the isolation regions 832 between the vertical NAND architecture NROM strings 804 can be extended into the substrate 808 to allow the formation of P-wells, where each P-well contains a single NAND string 804 and can be biased in isolation from the other strings 804 of the array 800. It is also noted that the control gates/word lines 806 and select gates 810/select lines 840 cross these isolation regions 832 so that each control gate/word address line 806 controls the operation of NROM memory cells 802 and each select line 840 the operation of select gates 810 across multiple NAND memory strings 804.

[0077] In Figures 8A-8C, the substrate 808 of the vertical NAND architecture NROM memory array 800 is P-doped. A substrate connection can be utilized that can allow for biasing of the P-doped substrate 808. It is noted that other forms of substrate doping, substrate biasing, and substrate types and regions (including, but not limited to silicon-on-sapphire (SOS) technology, silicon-on-insulator (SOI) technology, thin film transistor (TFT) technology, doped and undoped semiconductors, epitaxial layers of silicon supported by a base semiconductor) in embodiments of the present invention are possible and should be apparent to those skilled in the art with the benefit of the present invention.

[0078] It is noted that the NROM memory cells of embodiments of the present invention can be formed from multiple conventional materials. For the gate-insulator stack (gate insulator-trapping layer-top insulator) these materials may include, but are not limited to, oxide-nitride-oxide (ONO), oxide-nitride-aluminum oxide, oxide-aluminum oxide-oxide, oxide-silicon oxycarbide-oxide, composite layers of a silicon oxide-an oxide of Ti, Ta, Hf, Zr, or La, and a silicon oxide, and composite layers of an oxide-a non-stoichiometric oxide of Si, N, Al, Ti, Ta, Hf, Zr, and La, and an oxide. Additional trapping layer materials for NROM memory cell embodiments of the present invention, may also include, but are not limited to, wet oxides not annealed, silicon rich oxides, silicon rich aluminum oxide, silicon oxycarbide, silicon oxide with silicon

carbide nanoparticles, and non-stoichiometric oxides of Si, N, Al, Ti, Ta, Hf, Zr, and La.

[0079] As noted above the programming of the NROM memory cells of the vertical NAND and NOR NROM architecture memory structures, strings, and arrays of embodiments of the present invention can be accomplished by conventional tunnel injection of electrons by having a positive gate voltage with respect to the substrate or P-well. In another embodiment of the present invention, programming is accomplished by channel hot electron injection (HEI). Erasure of the NROM memory cells of embodiments of the present invention can be accomplished by conventional tunneling or negative voltages applied to the control gate voltages with respect to the substrate or P-well. With the above listed programming and erasure techniques, the NROM memory cells of embodiments of the present invention can be utilized for two-bit storage as a conventional planar NROM memory cells would be, storing charge in the trapping layer near each source/drain, allowing one bit to be read/programmed when biased in the forward direction and the other to be read/programmed when biased in the reverse direction.

[0080] In alternative embodiments of the present invention, substrate enhanced hot electron injection (SEHE) can be utilized for NROM memory cell programming and/or substrate enhanced band to band tunneling induced hot hole injection (SEBBHH) for NROM memory cell erasure. However, while the required voltages for these operations may be lower, they may only be suitable for single bit storage operation mode.

[0081] Figure 9 illustrates a functional block diagram of a memory device 900 that can incorporate the vertical NAND architecture NROM memory array 700 or vertical NOR architecture NROM memory cell array 600 of the present invention. The memory device 900 is coupled to a processor 910. The processor 910 may be a microprocessor or some other type of controlling circuitry. The memory device 900 and the processor 910 form part of an electronic system 920. The memory device 900 has been simplified to focus on features of the memory that are helpful in understanding the present invention.

[0082] The memory device includes an array of vertical NROM memory cells and select gates 930. In one embodiment, the memory cells are vertical NROM memory cells and the memory array 930 are arranged in banks of rows and columns. The control gates of each row of memory cells is coupled with a wordline while the drain and source connections of the memory cells are coupled to bitlines. As is well known in the art, the connection of the cells to the bitlines depends on whether the array is a NAND architecture or a NOR architecture.

[0083] An address buffer circuit 940 is provided to latch address signals provided on address/data bus 962. Address signals are received and decoded by a row decoder 944 and a column decoder 946 to access the memory array 930. It will be appreciated by those skilled in the art, with the benefit of the present description, that the size of address input on the address/data bus 962 depends on the density and architecture of the memory array 930. That is, the size of the input address increases with both increased memory cell counts and increased bank and block counts. It is noted that other address input manners, such as through a separate address bus, are also known and will be understood by those skilled in the art with the benefit of the present description.

[0084] The memory device 900 reads data in the memory array 930 by sensing voltage or current changes in the memory array columns using sense/buffer circuitry 950. The sense/buffer circuitry, in one embodiment, is coupled to read and latch a row of data from the memory array 930. Data input and output buffer circuitry 960 is included for bi-directional data communication over a plurality of data connections in the address/data bus 962 with the processor/controller 910. Write circuitry 955 is provided to write data to the memory array.

[0085] Control circuitry 970 decodes signals provided on control connections 972 from the processor 910. These signals are used to control the operations on the memory array 930, including data read, data write, and erase operations. The control circuitry 970 may be a state machine, a sequencer, or some other type of controller.

[0086] Since the vertical NROM memory cells of the present invention use a CMOS compatible process, the memory device 900 of Figure 9 may be an embedded device with a CMOS processor.

[0087] The memory device illustrated in Figure 9 has been simplified to facilitate a basic understanding of the features of the memory. A more detailed understanding of internal circuitry and functions of memories are known to those skilled in the art.

[0088] It is also noted that other vertical NAND and NOR architecture NROM memory strings, segments, arrays, and memory devices in accordance with embodiments of the present invention are possible and should be apparent to those skilled in the art with benefit of the present disclosure.

Conclusion

[0089] NROM EEPROM memory devices and arrays have been described that facilitate the use of vertical NROM memory cells and select gates in NOR or NAND high density memory architectures. Memory embodiments of the present invention utilize vertical select gates and NROM memory cells to form NOR and NAND NROM architecture memory cell strings, segments, and arrays. These NROM memory cell architectures allow for improved high density memory devices or arrays with integral select gates that can take advantage of the feature sizes semiconductor fabrication processes are generally capable of and yet do not suffer from charge separation issues in typical multi-bit NROM cells. The memory cell architectures also allow for mitigation of disturb and overerasure issues by placing the NROM memory cells behind select gates that isolate the memory cells from their associated bit/data lines and/or source lines.

[0090] Although specific embodiments have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that any arrangement, which is calculated to achieve the same purpose, may be substituted for the specific embodiment shown. This application is intended to cover any adaptations or variations of the present invention. Therefore, it is manifestly intended that this invention be limited only by the claims and the equivalents thereof.